# Some Fundamental Control Theory II: Feedback Linearization of Single Input Nonlinear Systems

**William J. Terrell**

**1. INTRODUCTION.** In Part I of this article [**12**] we characterized the single-input single-output linear

$$x' = Ax + bu(t) \tag{1a}$$

$$y = c^T x \tag{1b}$$

where $A$ is $n \times n$, $x$ and $b$ are $n \times 1$, and $c$ is $n \times 1$, that can be transformed by a nonsingular linear transformation, $z = Tx$, to a linear *companion form*

$$z' = Pz + du(t) \equiv \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & 1 \\ -k_n & -k_{n-1} & -k_{n-2} & \cdots & -k_1 \end{bmatrix} z + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} u, \tag{2}$$

which has the desirable property of *complete controllability*. By the use of state feedback, $u = Kx$ with $K$ a $1 \times n$ matrix, such systems may be expressed in the particularly simple form, $z_1^{(n)} = v$, where $v$ is a new reference input that is available for control purposes.

For convenience we summarize the main result of Part I in Theorem 1; see [**12**] for definitions of the relevant concepts.

**Theorem 1.** *System* (1a) *can be transformed by a nonsingular linear transformation, $z = Tx$, to the companion system* (2), *if and only if rank* $[b\, Ab \ldots A^{n-1}b] = n$ ((1a) *is completely controllable*). *When this is the case, $T$ is unique and*

$$T = \begin{bmatrix} \tau \\ \tau A \\ \vdots \\ \tau A^{n-1} \end{bmatrix}, \tag{3}$$

*where $\tau$ is the unique solution of*

$$\tau\,[b\, Ab \ldots A^{n-1}b] = [0 \ldots 0\; 1] = d^T. \tag{4}$$

*There exists a nonsingular transformation $z = Tx$ taking* (1a) *to the companion system* (2) *with $z_1 = y = c^T x$, if and only if*

$$rank \begin{bmatrix} c^T \\ c^T A \\ \vdots \\ c^T A^{n-1} \end{bmatrix} = n,$$

(*system* (1a), (1b) *is completely observable*) *and*

$$c^T[b\,Ab \dots A^{n-1}b] = [0 \dots 0\ 1] = d^T.$$

*When this is the case*, $T$ *is unique and is the matrix* (3) *with* $\tau = c^T$.

Our main goal in Part II is to generalize Theorem 1 for the simplest form of single input nonlinear systems. The exposition uses ideas from the theory of differential equations, linear algebra, and analysis, and basic concepts in differential geometry. The equivalence problem we consider is one of the fundamental results of geometric nonlinear control theory. The original formulation and solution of this equivalence problem in the single-input case is due to R. W. Brockett [**1**].

We begin with a calculation that recapitulates the essential calculations that established Theorem 1. It is also useful in the generalization of Theorem 1. Consider the following product of $n \times n$ matrices,

$$\begin{bmatrix} c^T \\ c^TA \\ \vdots \\ c^TA^{n-1} \end{bmatrix} [b\,Ab \cdots A^{n-1}b] = \begin{bmatrix} 0 & \cdots & \cdots & 0 & c^TA^{n-1}b \\ 0 & & & * & * \\ \vdots & & & & \vdots \\ 0 & * & & & \vdots \\ c^TA^{n-1}b & * & \cdots & \cdots & * \end{bmatrix}, \tag{5}$$

which results when $c$, $A$, $b$ satisfy $c^TA^kb = 0$ for $0 \le k \le n-2$. If $c^TA^{n-1}b \neq 0$ as well, then both matrices on the left must be nonsingular. In Part I we showed that, conversely, if one of the matrices on the left is nonsingular (for *some* $b$, respectively $c$), then the other matrix is *also* nonsingular (for some $c$, respectively $b$). Nonsingularity on the right in (5) implies an observability condition and a controllability condition [**12**]. The geometric interpretation of the zeros on the right hand side is that the null space of the linear functional $y = c^Tx$ is the $(n-1)$ dimensional space, *span*$\{b, Ab, \dots, A^{n-2}b\}$. The duality aspects of Part I arise from the duality between vectors and linear functionals in a finite dimensional vector space. Similar duality considerations in Part II involve the pairing of *vector fields* and *co-vector fields* (or *differential* 1-*forms*).

If we replace the linear vector field $Ax$ in (1) by $f(x)$, and the vector $b$ by a vector field $g(x)$, we are interested in determining exactly when the resulting single input nonlinear system $x' = f(x) + g(x)u$ may be transformed to the special form $y^{(n)} = v$ *by local coordinate change and state feedback*. If such a transformation is possible, one can exploit the special control-theoretic properties discussed in [**12**] for that special form. In such a case, a nonlinear system can be controlled using linear control methods.

**Example 1.** Suppose $x \in R^3$ and $u \in R$. Is the system

$$x' = f(x) + g(x)u \equiv \begin{bmatrix} \exp(-x_2) \\ x_1 \\ \frac{1}{2}x_1^2 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} u$$

equivalent to $y^{(3)} = v$? What is this equivalence, and how do we determine if such an equivalence is possible?

**2. LINEARIZATION AND FUNCTIONS OF RELATIVE DEGREE** $n$. Consider the single-input single-output nonlinear version of (1a) described by

$$x' = f(x) + g(x)u, \tag{6a}$$

where $f$ and $g$ are smooth vector fields defined in some open region $\mathscr{G}$ of $R^n$. The term *smooth* means that the components of $f$ and $g$ are continuously differentiable as often as required in our discussion. Equation (6a) may be augmented with an output,

$$y = h(x), \tag{6b}$$

where $h$ is a smooth real-valued function defined on $\mathscr{G}$. We are interested in necessary and sufficient conditions under which system (6a) is equivalent to a linear companion system. Such an equivalence, if possible, generally requires feedback in addition to coordinate change: our examples make it clear that (4), the condition that allows for a coordinate transformation *directly* to companion form in the linear case, cannot generally be expected here.

**2.1 Input-Output Linearization.** As in the linear case, we can try to use a known output $h$ to help define a local coordinate change, $z = T(x)$, and state feedback, $u = \alpha(x) + \beta(x)v$ (with $\beta(x) \neq 0$), in a neighborhood $U$ of $x_0$ in (6), to produce a linear input-output equation,

$$y^{(j)} = v. \tag{7}$$

In (7), $v$ is called the new *reference input*, and we would like to have $j = n$ in order to capture the full $n$-dimensional dynamical system. This is the Input-Output Linearization Problem (IOLP). The idea is that feedback by $u = \alpha(x)$ simplifies the system equations by cancelling nonlinearities when possible, and then subsequent controls $v$ are available to control the dynamics in a desired manner.

The next definition recalls the behavior of the special outputs that yield observability *and* a companion form in the linear case of [**12**]. The definition is motivated by the form of the terms that appear when a function is differentiated repeatedly along system (6a).

**Definition 1.** The Lie derivative of a real-valued function $h$ along a vector field $g$ is the real-valued function $L_g h$ defined by $L_g h(x) \equiv dh(x) \cdot g(x)$, where $dh(x)$ is the row gradient of $h$ at $x$. For iterated derivatives of this type we write $L_g^0 h = h$, and $L_g^k h = L_g(L_g^{k-1} h)$. The function $h$ has *relative degree* $j \geq 1$ with respect to (6a) at the point $x_0$ if

(i) $L_g L_f^k h(x) = 0$ for all $x$ in a neighborhood of $x_0$, for all $0 \leq k < j - 1$, and

(ii) $L_g L_f^{j-1} h(x_0) \neq 0$.

The relative degree is the number of differentiations of $h$ along the system that are needed to make $u$ appear explicitly. Thus, for relative degree $j = 1$, note that only condition (ii) is relevant. There may be points $x_0$ where a relative degree for a function is not defined. By continuity, conditions (i) and (ii) allow us to speak of a function $h$ having relative degree $j$ *in an open set* $U$ containing $x_0$.

**Example 2.** Consider the system

$$x_1' = \sin x_2 \tag{8}$$

$$x_2' = -x_1^2 + u. \tag{9}$$

The function $y = h(x) = x_2$ has relative degree 1 at all points since $L_g h(x) = 1$. On the other hand, $y = h(x) = x_1$ satisfies $L_g h = 0$, and $L_g L_f h(x) = L_g(\sin x_2) = \cos x_2$, so this function has relative degree 2 at the equilibrium $x_0 = 0$ of the unforced system.

Functions having relative degree $n$ are especially useful. Theorem 1 illustrates why linear functionals of relative degree $n$ were central in [**12**].

Suppose an output $y = h(x)$ has relative degree $n$ at $x_0$. Then $y$ and the first $n-1$ time derivatives of $y$ along (6a) yield the equations

$$\begin{bmatrix} y \\ y' \\ \vdots \\ y^{(n-1)} \end{bmatrix} = \begin{bmatrix} h(x) \\ L_f h(x) \\ \vdots \\ L_f^{n-1} h(x) \end{bmatrix} + \begin{bmatrix} 0 \\ \vdots \\ 0 \\ L_g L_f^{n-1} h(x) \end{bmatrix} u. \tag{10}$$

In Proposition 2 we show that if $h$ has relative degree $n$ at $x_0$, then the vector function $[h \; L_f h \; \cdots \; L_f^{n-1} h]^T$ has nonsingular Jacobian with respect to $x$ at $x_0$; thus, the functions $L_f^k h(x)$ for $0 \leq k \leq n-1$ can be used as component functions in a nonlinear coordinate transformation defined locally around $x_0$. The presence of the nonzero function $L_g L_f^{n-1} h(x)$ as a coefficient of $u$ prevents us from obtaining the linear, controllable companion form for $z = [y, y', \ldots, y^{(n-1)}]$ by coordinate transformation alone. However, the coordinate transformation *together with state feedback* does produce the simple form, $y^{(n)} = v$. We need only define

$$u = \frac{1}{L_g L_f^{n-1} h(x)} \left( -L_f^{n-1} h(x) + v \right), \tag{11}$$

where $v$ is the new reference input. This situation allows for control action on the nonlinear system by operating with the linear controllable form (7) (with $j = n$). Note that an assumption of relative degree $j < n$ leads to a linear input-output relation (7). Relative degree $n$ of $h$ at $x_0$ is a type of local observability condition: with $u \equiv 0$, the Inverse Function Theorem [**9**, p. 193] applied to (10) implies that $x$ is determined by $[y, y', \ldots, y^{(n-1)}]$ near $x_0$.

**2.2 Input-State Linearization.** If a relative degree $n$ function is not readily available as an output function, we still ask for a local coordinate change, $z = T(x)$, and feedback $u = \alpha(x) + \beta(x)v$ with $\beta(x) \neq 0$ near $x_0$, that produces a *controllable linear system* for $z$. This is often called the Input-State Linearization Problem (ISLP). As shown in [**12**], given a controllable linear system, an additional coordinate change and state feedback may be used to produce the linear system

$$z' = Nz + dv \equiv \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix} z + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} u, \tag{12}$$

where $N$ is the standard nilpotent block with ones on the superdiagonal and zeros elsewhere, and $d = [0 \ldots 0 \; 1]^T$ as usual. Thus we take the ISLP to mean the problem of achieving the form (12).

Our first goal is to show that the ISLP is solvable if and only if there exists a function of relative degree $n$ at $x_0$ for (6a). We have already discussed the sufficiency of the relative degree $n$ condition, though the proof of Proposition 2 is needed to complete that discussion. After the next example, we show that solvability of the ISLP implies the existence of a function $\lambda(x)$ of relative degree $n$ at $x_0$ —namely, the first coordinate function $T_1$ of the transformation $T$. Our ultimate goal is to derive computable geometric conditions for the existence of a relative degree $n$ function.

One can adopt the point of view that there is no essential mathematical difference between the linearization problems IOLP and ISLP when a relative degree $n$ function exists—the difference lies in whether there is an *explicitly known* output function of relative degree $n$, or if such a function must be *constructed.* From the practical point of view, of course, this is an essential difference. We return to Example 2 to illustrate an important consideration.

**Example 3.** Consider again the system,

$$x_1' = \sin x_2$$

$$x_2' = -x_1^2 + u.$$

It is easy to linearize the input-output behavior; take $y = x_2$, and set $u = x_1^2 + v$, which produces the linear input-output relation $y' = v$. This relation would make it easy to track a *specified* output $y(t) = x_2(t)$ by control $v$, but the linearizing control $u$ would make $x_1$ unobservable if it is only the resulting input-output relation $y' = v$ that is considered. Thus, some of the original two-dimensional dynamics is hidden by the input-output relation. The presence of such unobservable dynamics introduces the question of internal stability for that dynamics; the $x_1$ variable may not remain well-behaved (e.g., bounded) when using a control strategy based on the relation $y' = v$. For example, suppose we wanted to hold the output $y = x_2$ via feedback at a constant value, $x_2 = c$. Then the $x_1$ solution would be $x_1(t) = x_1(0) + t\sin c$, and therefore $x_1 \to \infty$ as $t \to \infty$.

We now show that a solution of the ISLP entails a relative degree $n$ function. Suppose the ISLP is solved, so that the transformation $z = T(x)$, combined with feedback $u = \alpha(x) + \beta(x)v$, produces the linear system (12). From the definitions of the variables, this occurs if and only if

$$\frac{\partial T}{\partial x}(f(x) + g(x)u) = NT(x) + dv, \tag{13}$$

where equality holds *for all* $u = \alpha(x) + \beta(x)v$, with $v$ arbitrary. By considering $u = 0$ and $u = 1$, we see that (13) is equivalent to the two partial differential equations:

$$\frac{\partial T}{\partial x}f(x) = NT(x) - d\alpha(x)\beta^{-1}(x), \tag{14a}$$

$$\frac{\partial T}{\partial x}g(x) = d\beta^{-1}(x). \tag{14b}$$

If $T(x) = [T_1(x)\, T_2(x) \dots T_n(x)]^T$, then

$$NT(x) - d\alpha(x)\beta^{-1}(x) = \left[T_2(x)\, T_3(x) \dots T_n(x) - \alpha(x)\beta^{-1}(x)\right]^T,$$

and

$$d\beta^{-1}(x) = \left[0\,0 \dots 0\beta^{-1}(x)\right]^T.$$

This allows us to display (14a) in detail:

$$dT_k(x)\cdot f(x) = L_f T_k(x) = T_{k+1}(x), \quad k = 1, \dots, n-1, \tag{15a}$$

and

$$dT_n(x)\cdot f(x) = -\alpha(x)\beta^{-1}(x). \tag{15b}$$

Therefore all components of $T$ are determined by the *first* component $T_1$. And the last equation specifies the ratio, $-\alpha(x)\beta^{-1}(x)$, in terms of $T_1$. Similarly we display (14b) using (15a) and Definition 1:

$$dT_{k+1}(x)\cdot g(x) = L_g L_f^k T_1(x) = 0, \quad k = 0,\dots,n-2, \tag{16a}$$

while

$$dT_n(x)\cdot g(x) = L_g L_f^{n-1} T_1(x) = \beta^{-1}(x) \neq 0. \tag{16b}$$

Thus, by (15a) and (16), *the solution of the ISLP requires a function $T_1(x)$ that has relative degree $n$*. Given $T_1$, the rest of $T$ is then defined by (15a), and the required feedback $u = \alpha(x) + \beta(x)\,v$ is determined by

$$\beta(x) = \frac{1}{L_g L_f^{n-1} T_1(x)}; \quad \alpha(x) = -\frac{L_f^{n-1} T_1(x)}{L_g L_f^{n-1} T_1(x)}. \tag{17}$$

Notice that an equilibrium point of interest, say $x_0$ such that $f(x_0) = 0$, can always be transformed to $z = T(x_0) = 0$: by (15a), we need only require that $T_1(x_0) = 0$. We summarize the preceding discussion as follows.

**Theorem 2.** *System* (6a) *can be transformed by coordinate transformation $z = T(x)$ and state feedback $u = \alpha(x) + \beta(x)v$ in a neighborhood of $x_0$ to the linear controllable form $z' = Nz + dv$ if and only if there exists a function $\lambda$ having relative degree $n$ with respect to* (6a) *at the point $x_0$. When this is the case, $T$ is determined by* (15a) *by defining the first component function by $T_1 = \lambda$, and the feedback is determined by* (17).

*Proof:* Everything has been proved except the legitimacy of the coordinate change

$$z = T(x) = \left[\lambda(x), L_f\lambda(x),\dots, L_f^{n-1}\lambda(x)\right]^T,$$

when $\lambda$ has relative degree $n$ at $x_0$. This argument appears as part of a more general statement in Proposition 2. Assuming this result, Theorem 2 follows. ■

Let us illustrate these ideas, and the general *nonuniqueness* of $T$ in the nonlinear case:

**Example 4.** Consider the system in Example 3 near the equilibrium $x_0 = 0$ of the unforced system (with $u = 0$). The equations (16) for $T_1$ are

$$L_g T_1(x) = 0, \quad L_g T_2 \neq 0; \quad T_2 = dT_1(x)\cdot f(x).$$

Using $g = [0\;1]^T$ we see that $T_1$ is independent of $x_2$, so $T_2(x) = (\partial T_1/\partial x_1)\sin x_2$. The nontriviality condition becomes

$$L_g T_2 = \frac{\partial T_2}{\partial x_2} = \frac{\partial T_1}{\partial x_1}\cos x_2 \neq 0,$$

which holds as long as $\cos x_2 \neq 0$ and $\partial T_1/\partial x_1 \neq 0$. One solution is $T_1(x) = x_1$, which we used in Example 3. The choice is not unique: $T_1(x) = x_1 - x_1^2$ also works near $x_0 = 0$ in fulfilling the conditions of Theorem 2.

**Example 5.** Using Theorem 2 we can show that Example 1 is not input-state linearizable. The vector field $g$ in Example 1 is $g = [1\;\;0\;\;0]^T$. Let $\lambda$ be a smooth function, and suppose $L_g\lambda(x) = (\partial\lambda/\partial x_1)(x) = 0$ for $x$ in an open set $U$; then $\lambda$

is independent of $x_1$ in $U$, so $\lambda = \lambda(x_2, x_3)$ there. If we also impose the condition that $L_g L_f \lambda(x) = 0$ in $U$, then we have

$$L_g(d\lambda \cdot f)(x) = \frac{\partial}{\partial x_1}\left(\frac{\partial \lambda}{\partial x_2} x_1 + \frac{\partial \lambda}{\partial x_3}\frac{1}{2}x_1^2\right)$$
$$= \frac{\partial \lambda}{\partial x_2} + x_1 \frac{\partial \lambda}{\partial x_3} = 0;$$

and it is not possible for this to hold in $U$ when $\lambda = \lambda(x_2, x_3)$. Thus, there is no function having relative degree 3 with respect to this system in any open set in $R^3$.

**3. NECESSARY CONDITIONS FOR ISLP SOLVABILITY.** We've seen that an "artificial output" $T_1$ of relative degree $n$ is determined by a *nontrivial* solution of the partial differential equation system (16a), the nontriviality condition being (16b). When is this first order partial differential equation system solvable? We want computable conditions for solvability directly in terms of $f$ and $g$. Let us first consider necessary conditions, and for that, a nonlinear version of the calculation in (5) is helpful, with an appropriate generalization of the columns $b$, $Ab, \ldots$ in (5). This analysis of necessary conditions completes the proof of Theorem 2 and leads to the computable criteria we seek. We should remark that there are indeed local observability and controllability conditions involved in what follows, as you might expect. After all, we are transforming to a linear system with the properties of controllability and observability. We discuss a local *reachability* property after the main Theorem 3—this property is related to, but weaker than, complete controllability. For now, we continue to focus on the goal of generalizing the transformation to companion form statements of Theorem 1, but in the process we indicate the intuition involved in generalizations of the linear controllability criterion.

Let us consider (5): we replace the first matrix on the left by $[d\lambda \, dL_f\lambda \ldots dL_f^{n-1}\lambda]^T$, where $\lambda$ has relative degree $n$, and we put $g(x)$ in place of $b$ in the second matrix on the left; we then need vector fields to replace $Ab, \ldots, A^{n-1}b$. That is, we need to identify the null space of the differential $d\lambda$. The appropriate vector fields can be motivated either analytically or algebraically, and we consider both aspects in order to build some intuition.

The next definition can be motivated by the calculation of Proposition 1, but it is convenient to state it here and follow it with an important example.

**Definition 2.** The Lie bracket $[g_1, g_2]$ of two vector fields $g_1$, $g_2$ is the vector field defined by

$$[g_1, g_2](x) \equiv \frac{\partial g_2}{\partial x}(x) g_1(x) - \frac{\partial g_1}{\partial x}(x) g_2(x). \tag{18}$$

Here is a notation that helps with iterated brackets: define $ad_{g_1}^0 g_2 = g_2$, $ad_{g_1} g_2 = [g_1, g_2]$, and $ad_{g_1}^k g_2 \equiv [g_1, ad_{g_1}^{k-1} g_2]$ for $k \geq 1$. The brackets described in (18) are important in the linear system case:

**Example 6.** If $f(x) = Ax$ and $g(x) = b$, then $[f, g](x) = -Ab$. Also, $ad_f^2 g(x) = [f,[f, g]](x) = A^2 b$, and in general, $ad_f^k g(x) = (-1)^k A^k b$.

Example 6 suggests that the brackets $ad_f^k g(x)$ are important in the nonlinear case. To confirm this, we consider an analytic relation involving $f$ and $g$. Write

$\phi_t^\eta(x)$ for the time $t$ solution map of the differential equation $x' = \eta(x)$; that is, $\phi_0^\eta(x) = x$ and $\partial\phi_t^\eta(x)/\partial t = \eta(\phi_t^\eta(x))$. Then $(\phi_t^\eta)^{-1} = \phi_{-t}^\eta$ where defined. Also write $(\phi_t^\eta)_*$ for the derivative map with Jacobian matrix $\partial\phi_t^\eta/\partial x$. The *variational differential equation* satisfied by $(\phi_t^\eta)_*$ is the equation

$$\frac{\partial}{\partial t}(\phi_t^\eta)_* = \frac{\partial\eta}{\partial x}(\phi_t^\eta)_*;$$

to prove this relation, one differentiates the identity $d/dt(\phi_t^\eta(x)) = \eta(\phi_t^\eta(x))$ with respect to $x$ and then interchanges the order of differentiations with respect to $x$ and $t$. It is also useful to have the variational equation for $(\phi_{-t}^\eta)_* = (\phi_t^\eta)_*^{-1}$: the formula

$$\frac{d}{dt}A^{-1}(t) = -A^{-1}(t)\frac{dA}{dt}A^{-1}(t)$$

for the derivative of the inverse of a nonsingular matrix function of $t$, applied to $A = (\phi_t^\eta)_*$, gives

$$\frac{\partial}{\partial t}(\phi_{-t}^\eta)_* = -(\phi_{-t}^\eta)_*\frac{\partial\eta}{\partial x}.$$

Now consider following the flow of $f$ (the vector field in (1a)) for a short time $t$, to the point $\phi_t^f(x_0)$; then determine the direction vector $(f+g)(\phi_t^f(x_0))$, the direction you would move if you turned the control $u$ "on" with $u = 1$; and, finally, transfer this tangent vector back to the point $x_0$ by applying $(\phi_{-t}^f)_*$. Thus, consider the "curve of tangents" based at $x_0$,

$$V(t) = (\phi_{-t}^f)_*\eta(\phi_t^f(x_0)), \tag{19}$$

where $\eta = f + g$, and more specifically, the derivative $V'(0)$ at $t = 0$. By considering shorter and shorter times $t$, it is plausible that the vector $V'(0)$ (and also the higher order derivatives of $V$ at $t = 0$) should indicate something about the directions we might move, starting at $x_0$, by some suitable "off-on" switching strategy for the input $u$. This can be made precise, in a way that provides an alternative motivation for the Lie bracket operation; see [**8**, Proposition 3.6, pp. 77–78] or [**7**, pp. 323–324]. We have motivated (19) here because it is useful in the proof of the main Theorem 3. Notice that the construction in (19) is valid for *any* vector field $\eta$, although $\eta = f + g$ is the immediate interest.

**Proposition 1.** *The tangent vector $V'(0)$ of the curve* (19) (*where $\eta = f + g$*) *is*

$$V'(0) = [f,\eta](x_0) = [f,g](x_0).$$

*If $f$ and $g$ are analytic, then $V^{(k)}(0) = ad_f^k g(x_0)$ for all $k \geq 1$.*

*Proof:* Use the formula for the derivative of a product, together with the variational equations, to compute

$$\begin{aligned} V'(t) &= \frac{\partial}{\partial t}(\phi_{-t}^f)_*\eta(\phi_t^f(x_0)) + (\phi_{-t}^f)_*\frac{\partial\eta}{\partial x}\frac{\partial}{\partial t}(\phi_t^f(x_0)) \\ &= -(\phi_{-t}^f)_*\frac{\partial f}{\partial x}\eta(\phi_t^f(x_0)) + (\phi_{-t}^f)_*\frac{\partial\eta}{\partial x}f(\phi_t^f(x_0)) \\ &= (\phi_{-t}^f)_*[f,\eta](\phi_t^f(x_0)). \end{aligned} \tag{20}$$

Set $t = 0$ to get $V'(0) = [f,\eta](x_0)$, taking into account the initial conditions $\phi_0^f(x) = x$ and $(\phi_0^f)_* = I$. Since $[f, f+g] = [f,g]$, the first statement is proved. Given (20), the second statement follows by induction. ■

Because of Proposition 1, the Lie bracket $[f, \eta]$ is also called the *Lie derivative of $\eta$ along $f$*.

The algebraic property that shows the connection between the Lie bracket operation on vector fields and the Lie derivative operation on functions is the *Jacobi identity*; it says that if $v$, $w$ are vector fields and $\lambda$ is a smooth function, then $(L_{[v,w]}\lambda)(x) = (L_v L_w \lambda - L_w L_v \lambda)(x)$. This identity is proved as follows. For any smooth $\lambda$, and any $x$,

$$L_v L_w \lambda(x) - L_w L_v \lambda(x) = L_v(d\lambda(x) \cdot w(x)) - L_w(d\lambda(x) \cdot v(x))$$
$$= \left( d^2\lambda(x) \cdot w(x) + d\lambda(x) \cdot \frac{\partial w}{\partial x}(x) \right) \cdot v(x)$$
$$- \left( d^2\lambda(x) \cdot v(x) + d\lambda(x) \cdot \frac{\partial v}{\partial x}(x) \right) \cdot w(x), \quad (21)$$

where $d^2\lambda(x)$ is the matrix of second partial derivatives of $\lambda$ at $x$. Since $d^2\lambda$ is symmetric, we get

$$L_v L_w \lambda - L_w L_v \lambda = d\lambda \cdot \left( \frac{\partial w}{\partial x} v - \frac{\partial v}{\partial x} w \right) = d\lambda \cdot [v, w] = L_{ad_v w} \lambda.$$

The Jacobi identity itself helps to identify $\ker d\lambda(x)$ for $x \in U$, for a function $\lambda$ having relative degree $n$. Given $g$ as the replacement for $b$ in (5), and assuming that $d\lambda(x) \cdot g(x) = 0$, one can show that $ad_f^k g(x) \in \ker d\lambda(x)$ for $k = 1, \ldots n-2$ by induction, using the Jacobi identity. We now give the details of the nonlinear version of the calculation in (5).

**Proposition 2.** *If $\lambda$ has relative degree $n$ with respect to* (6a) *in the open set $U$, then for all $x \in U$,*

(1) *the covectors $d\lambda(x)$, $dL_f\lambda(x), \ldots, dL_f^{n-1}\lambda(x)$ are linearly independent*;
(2) *the vectors $g(x)$, $ad_f g(x), \ldots, ad_f^{n-1} g(x)$ are linearly independent.*

*Proof:* Consider the matrix product that generalizes (5):

$$\begin{bmatrix} d\lambda(x) \\ dL_f\lambda(x) \\ \vdots \\ dL_f^{n-1}\lambda(x) \end{bmatrix} \begin{bmatrix} g(x) & ad_f g(x) \ldots ad_f^{n-1} g(x) \end{bmatrix}$$

$$= \begin{bmatrix} L_g\lambda(x) & L_{ad_f g}\lambda(x) & \cdots & \cdots & L_{ad_f^{n-1} g}\lambda(x) \\ L_g L_f \lambda(x) & & & L_{ad_f^{n-2} g} L_f \lambda(x) & \star \\ \vdots & & & & \vdots \\ L_g L_f^{n-1}\lambda(x) & \star & \cdots & \cdots & \star \end{bmatrix}. \quad (22)$$

We now use the relative degree $n$ assumption and the Jacobi identity to show that the matrix on the right in (22) is lower right triangular with nonzero entries on the skew-diagonal; Proposition 2 then follows.

By relative degree $n$, the first column has the required form. Proceed by induction on the columns, using the Jacobi identity. Now the $k, l$ entry in the matrix is $(dL_f^k h) \cdot (ad_f^l g)$, for $0 \le k, l \le n-1$. The diagonal entries in question are those for which $k + l = n - 1$. Assume the desired property for column $l$: thus, assume that $L_{ad_f^l g} L_f^k h = 0$ for $k + l \le n - 2$, and $L_{ad_f^l g} L_f^{n-1-l} h \ne 0$. For column $l + 1$ we need $L_{ad_f^{l+1} g} L_f^k h = 0$ for $k \le n - 3 - l$, and $L_{ad_f^{l+1} g} L_f^{n-2-l} h \ne 0$. Using the Jacobi identity, column $l + 1$ is, for $0 \le k \le n - 1$:

$$(k, l)\ entry = \left(dL_f^k h\right) \cdot \left(ad_f^{l+1} g\right) = L_f L_{ad_f^l g} L_f^k h - L_{ad_f^l g} L_f^{k+1} h. \tag{23}$$

Apply the hypothesis to (23) for $k = 0, \dots, n - 3 - l$ to get zero. For $k = n - 2 - l$, only the last term in (23) contributes to the skew-diagonal entry, which is

$$-L_{ad_f^l g} L_f^{n-1-l} h \ne 0.$$

Notice that this is the *negative* of the skew-diagonal entry in column $l$. Since the first column has last entry $L_g L_f^{n-1} h(x)$, the diagonal entry in column $l$ must therefore be $(-1)^l L_g L_f^{n-1} h(x) \ne 0$ for $l = 0, \dots, n - 1$. Thus, our matrix is lower right triangular with nonzero skew-diagonal entries for $x \in U$. ∎

Notice that Proposition 2 completes the proof of Theorem 2, because it shows that if $\lambda$ has relative degree $n$ then the vector function

$$\left[\lambda(x), L_f \lambda(x), \dots, L_f^{n-1} \lambda(x)\right]^T$$

has a nonsingular Jacobian for $x \in U$.

As in (5), a geometric interpretation of Proposition 2 is that the null space of the differential $d\lambda(x)$ is the $(n - 1)$-dimensional space

$$\mathscr{D}(x) = span\left\{g(x), ad_f g(x), \dots, ad_f^{n-2} g(x)\right\}, \quad x \in U. \tag{24}$$

We view each subspace $\mathscr{D}(x)$ as a subspace of the tangent space of $R^n$ at $x$, $T_x R^n \approx R^n$. The collection of the subspaces (24) for $x \in U$ is called a *distribution* on $U$. In the linear case, a constant distribution such as $span\{b, Ab, \dots, A^{n-2} b\}$ is automatically the null space of a nonzero linear functional. Proposition 2 shows that the nonsingularity of both factors on the left in (22) is *necessary* for the solution of the ISLP. Under the assumption of Proposition 2, there is an additional necessary condition on the distribution (24), which is not revealed in the calculations of Proposition 2. The additional condition on $\mathscr{D}(x)$ is the geometric condition of *involutivity.* Involutivity is an integrability condition that guarantees that the distribution (24) is the space annihilated by the differential of a function having relative degree $n$. The next section examines this concept, and develops the geometric conditions for the solvability of the ISLP.

## 4. GEOMETRIC CRITERIA FOR ISLP SOLVABILITY.

It's convenient to place some formal definitions here.

**Definition 3.** A distribution $\mathscr{D}$ on $U$ is a smooth assignment (via functions like $g$, $ad_f g$, etc. in (24)) of a subspace of the tangent space $T_x U \approx R^n$, for $x \in U$. A distribution $\mathscr{D}(x)$ is *involutive* in $U$ if, for vector fields $v_1$ and $v_2$, and all $x \in U$,

$$v_1(x), v_2(x) \in \mathscr{D}(x) \Rightarrow [v_1, v_2](x) \in \mathscr{D}(x).$$

A distribution $\mathscr{D}$ is *nonsingular* in $U$ if dim $\mathscr{D}(x)$ is constant in $U$. A nonsingular distribution $\mathscr{D}$ with dim $\mathscr{D}(x) = k$ is *integrable* in $U$ if there are $n - k$ functions $\lambda_j$ such that $span\,\{d\lambda_j(x) : 1 \le j \le n - k\} = \mathscr{D}^\perp(x)$, or equivalently, $\bigcap_{i=1}^{n-k} \ker d\lambda_j(x) = \mathscr{D}(x)$.

To illustrate the involutivity concept, we return to Example 1.

**Example 7.** The distribution (24) for Example 1 is $\mathscr{D}(x) = span\{g(x), ad_f g(x)\}$. Appropriate bracket calculations for Example 1 give

$$ad_f g(x) = -\begin{bmatrix} 0 & -e^{-x_2} & 0 \\ 1 & 0 & 0 \\ x_1 & 0 & 0 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ x_1 \end{bmatrix},$$

and then

$$[g, ad_f g](x) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

If we form the matrix

$$\begin{bmatrix} g(x) & ad_f g(x) & [g, ad_f g](x) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & x_1 & 1 \end{bmatrix},$$

we see that its rank is everywhere equal to 3, so that $[g, ad_f g](x)$ does not lie in $\mathscr{D}(x)$ for any $x$. It follows that the distribution $\mathscr{D} = span\{g, ad_f g\}$ is not involutive in any open set. Compare this situation with Example 5.

*Frobenius' Theorem* states that for a nonsingular distribution $\mathscr{D}$ (of *any* dimension), integrability is *equivalent* to involutivity [**4**, p. 23]. The following discussion of the ideas of Frobenius' Theorem in the case of interest for the Input-State Linearization Problem uses virtually all the tools discussed so far.

It is easy to show that involutivity is *necessary* for $\mathscr{D}(x)$ in (24) to be the tangent space at $x$ of the level set within $U$ of a smooth function $\lambda$ having relative degree $n$. If $\lambda$ has relative degree $n$ at $x_0$, Proposition 2 shows that $[g(x)\, ad_f g(x) \dots ad_f^{n-1} g(x)]$ has rank $n$ for all $x \in U$, so the distribution $\mathscr{D}$ is nonsingular with constant dimension $n - 1$ in $U$. Moreover, we have $d\lambda(x)[g(x)\, ad_f g(x) \dots ad_f^{n-2} g(x)] = 0$. Thus, the pointwise orthogonal complement of $\mathscr{D}$ in $U$ is $\mathscr{D}^{\perp} = span\{d\lambda\}$. This says that the distribution $\mathscr{D}$ is integrable on $U$. We now show involutivity directly. If $0 \leq i, j \leq n - 2$, then for $x \in U$,

$$d\lambda(x) \cdot [ad_f^i g, ad_f^j g](x) = L_{ad_f^i g} L_{ad_f^j g} \lambda(x) - L_{ad_f^j g} L_{ad_f^i g} \lambda(x) = 0$$

since $\lambda$ has relative degree $n$. Therefore $[ad_f^i g, ad_f^j g](x) \in \mathscr{D}(x)$ for all $x$ in $U$. This is sufficient to show that $\mathscr{D}$ is involutive, since any two vector fields in $\mathscr{D}$ are linear combinations of the ones we just dealt with, and one can show that for fields $\xi$, $\eta$ in $\mathscr{D}$ and functions $a$, $b$, we have

$$\begin{aligned} [a\xi, b\eta](x) = a(x)b(x)[\xi, \eta](x) &+ (L_\xi b(x))a(x)\eta(x) \\ &- (L_\eta a(x))b(x)\xi(x). \end{aligned} \tag{25}$$

To establish (25), use the Jacobi identity plus the characterization that vector fields $v$, $w$ are equal if and only if $L_v \lambda = L_w \lambda$ for all smooth functions $\lambda$.

Theorem 3 provides the goal of computable geometric criteria on $f$ and $g$ for the solution of the Input-State Linearization Problem. Involutivity is the necessary *and* sufficient integrability condition for the system of partial differential equations (16a).

**Theorem 3.** *The system $x' = f(x) + g(x)u$ is input-state linearizable in a neighborhood $U$ of $x_0$ if and only if*

(I) $[g(x_0)\, ad_f g(x_0) \dots ad_f^{n-1} g(x_0)]$ *has rank $n$, and*
(II) *the distribution $\mathscr{D} = span\{g, ad_f g, \dots, ad_f^{n-2} g\}$ is involutive in $U$.*

*Proof:* We show that (I) and (II) are equivalent to the existence of a function $\lambda$ having relative degree $n$ at $x_0$. *Necessity* of (I) is covered by Proposition 2, and necessity of (II) was discussed before the theorem statement.

*Sufficiency.* If (I) and (II) hold, then the distribution $\mathscr{D}$ is nonsingular on $U$ with dimension $n - 1$. Clearly, there is a smooth covector field in $U$ defined by a smooth row vector $w(x)$ such that $w(x) = span\, \mathscr{D}^{\perp}(x)$ and also $w(x) \cdot ad_f^{n-1} g(x) \neq 0$ for $x \in U$. Thus, for all $x \in U$, we have

$$span\left\{g(x), \dots, ad_f^{n-2} g(x), w(x)\right\} = R^n.$$

The function $\lambda$ can be constructed from the flows of these vector fields. To simplify notation for this argument, define $v_1 = g$, $v_2 = ad_f g, \dots, v_{n-1} = ad_f^{n-2} g$, $v_n = w$.

Let $U_\epsilon$ be a ball of radius $\epsilon$ about the zero vector in $R^n$. There is an $\epsilon > 0$ such that the map $\psi : U_\epsilon \to \psi(U_\epsilon) \subset U$ defined by

$$\psi(z) = \psi(z_1, \dots, z_n) = \phi_{z_1}^{v_1} \circ \phi_{z_2}^{v_2} \circ \cdots \circ \phi_{z_n}^{v_n}(x_0)$$

is a diffeomorphism onto its image, that is, $\psi$ is smooth, one-to-one, and has a smooth inverse map defined on $\psi(U_\epsilon)$. This is because repeated application of the chain rule shows that at $z = 0$ we have

$$\frac{\partial \psi}{\partial z_i}(0) = v_i(x_0), \quad i = 1, \dots, n, \tag{26}$$

and by hypothesis, the vectors $v_i(x_0)$ are independent. Thus, the Inverse Function Theorem ensures that there is an $\epsilon > 0$ so that $\psi$ is a local diffeomorphism onto its image. The $z$ coordinates are time coordinates that "straighten out" the flows for the vector fields $v_i$. Write the inverse of $\psi$ in the form

$$\psi^{-1}(x) = \begin{bmatrix} \lambda_1(x) \\ \vdots \\ \lambda_n(x) \end{bmatrix}.$$

Now consider (26) together with the identity

$$\left(\frac{\partial \psi^{-1}}{\partial x}\right)_{z=\psi^{-1}(x)} \left(\frac{\partial \psi}{\partial z}\right)_{x=\psi(z)} = I_{n \times n}. \tag{27}$$

The strategy is to show that $d\lambda_n(x)$ spans $\mathscr{D}^{\perp}(x)$, by showing that the first $n-1$ columns of $(\partial\psi/\partial z)_{x=\psi(z)}$ form a basis of $\mathscr{D}(x)$ at any $x \in U$, for then (27) implies that $span\{d\lambda_n(x)\} = \mathscr{D}^{\perp}(x)$ for $x \in U$.

Using the chain rule, we find that the $i$-th column of $\partial\psi/\partial z$ is

$$\begin{aligned} \frac{\partial \psi}{\partial z_i} &= \left(\phi_{z_1}^{v_1}\right)_* \circ \cdots \circ \left(\phi_{z_{i-1}}^{v_{i-1}}\right)_* \frac{\partial}{\partial z_i}\left(\phi_{z_i}^{v_i} \circ \cdots \circ \phi_{z_n}^{v_n}(x_0)\right) \\ &= \left(\phi_{z_1}^{v_1}\right)_* \circ \cdots \circ \left(\phi_{z_{i-1}}^{v_{i-1}}\right)_* v_i\left(\phi_{z_i}^{v_i} \circ \cdots \circ \phi_{z_n}^{v_n}(x_0)\right) \\ &= \left(\phi_{z_1}^{v_1}\right)_* \circ \cdots \circ \left(\phi_{z_{i-1}}^{v_{i-1}}\right)_* v_i\left(\phi_{-z_{i-1}}^{v_{i-1}} \circ \cdots \circ \phi_{-z_1}^{v_1}(\psi(z))\right), \end{aligned}$$

where we use the fact that $(\phi_z^{v_i})^{-1} = \phi_{-z}^{v_i}$ for the local flows of the $v_i$. If we show that, whenever $\xi$, $\eta$ are vector fields that are pointwise in $\mathscr{D}$ we also have

$$(\phi_t^{\eta})_* \xi(\phi_{-t}^{\eta}(x)) \in \mathscr{D}(x),$$

then it follows that the column vector $(\partial\psi/\partial z_i)_{z=\psi^{-1}(x)}$ is also in $\mathscr{D}(x)$. Since any vector field $\xi$ in $\mathscr{D}$ can be written as $\Sigma c_i v_i$ for some functions $c_i$ in $\mathscr{D}$, we need only consider the case when $\xi = v_i$ for some $i = 1, \ldots, n-1$.

Thus, for a vector field $\eta \in \mathscr{D}$, let $x$ be *a fixed point* in $\mathscr{D}$, and set

$$V_i(t) = (\phi_{-t}^{\eta})_* v_i(\phi_t^{\eta}(x)); \quad i = 1, \ldots, n-1. \tag{28}$$

The vector functions $V_i$ are defined for some interval of $t$ about 0. It follows from Proposition 1 that

$$\frac{d}{dt} V_i(t) = (\phi_{-t}^{\eta})_* [\eta, v_i](\phi_t^{\eta}(x)).$$

Since $\mathscr{D}$ is involutive and $\eta, v_i \in \mathscr{D}$, there exist functions $\alpha_{ij}$ defined around $x$ such that

$$[\eta, v_i] = \sum_{j=1}^{n-1} \alpha_{ij} v_j,$$

so that

$$\frac{d}{dt} V_i(t) = (\phi_{-t}^{\eta})_* \left( \sum_{j=1}^{n-1} \alpha_{ij} v_j \right) (\phi_t^{\eta}(x)) = \sum_{j=1}^{n-1} \alpha_{ij}(\phi_t^{\eta}(x)) V_j(t) .$$

Thus, $V = [V_1 \ldots V_{n-1}]$ is a matrix solution of a linear system of differential equations having the form $V' = VA^T$, where $A \equiv [\alpha_{ij}]$ and $1 \le i, j \le n-1$. Therefore we can write

$$[V_1(t) \ldots V_{n-1}(t)] = [V_1(0) \ldots V_{n-1}(0)] X(t), \tag{29}$$

where $X(t)$ is an $(n-1) \times (n-1)$ fundamental matrix of solutions. Multiply (29) from the left by $(\phi_t^{\eta})_*$ and use (28) to get

$$[v_1(\phi_t^{\eta}(x)) \ldots v_{n-1}(\phi_t^{\eta}(x))] = [(\phi_t^{\eta})_* v_1(x) \ldots (\phi_t^{\eta})_* v_{n-1}(x)] X(t).$$

Now, for small $t$ we may replace $x$ by $\phi_{-t}^{\eta}(x)$ on the same orbit to get

$$[v_1(x) \ldots v_{n-1}(x)] = [(\phi_t^{\eta})_* v_1(\phi_{-t}^{\eta}(x)) \ldots (\phi_t^{\eta})_* v_{n-1}(\phi_{-t}^{\eta}(x))] X(t).$$

Since $X(t)$ is nonsingular, we get for $i = 1, \ldots, n-1$,

$$\begin{aligned}(\phi_t^{\eta})_* v_i(\phi_{-t}^{\eta}(x)) &\in span\{v_1(x), \ldots, v_{n-1}(x)\} \\ &= span\{g(x), \ldots, ad_f^{n-2} g(x)\} = \mathscr{D}(x).\end{aligned}$$

The proof that the columns $\partial\psi/\partial z_i$ are in $\mathscr{D}$ for $i = 1, \ldots, n-1$ is now complete.

It remains to show that $\lambda_n$ satisfies the nontriviality condition required for relative degree $n$. But from the identity (27) and the conditions (26) we have

$$d\lambda_n(x_0) \cdot \frac{\partial\psi}{\partial z_n}(0) = d\lambda_n(x_0) \cdot v_n(x_0) = d\lambda_n(x_0) \cdot w(x_0) = 1.$$

Thus, $d\lambda_n(x_0)$ is parallel to $w(x_0) \in \mathscr{D}^{\perp}(x_0)$, so $d\lambda_n(x_0) \cdot ad_f^{n-1} g(x_0) \neq 0$. Now apply the Jacobi identity to

$$L_{ad_f^{n-1} g} \lambda_n(x_0) = L_{[f, ad_f^{n-2} g]} \lambda_n(x_0),$$

to conclude that $-L_{ad_f^{n-2}g} L_f \lambda_n(x_0) = (-1)^{n-1} L_g L_f^{n-1} \lambda(x_0) \neq 0$. Therefore $\lambda_n$ has relative degree $n$ in some open set in $U$. ■

The proof of sufficiency in Theorem 3 follows the proof in [**4**, pp. 24–28] that involutivity implies integrability. We really needed a particular *codimension one* case of Frobenius' Theorem, where dim $\mathscr{D}^{\perp} = 1$ and $\mathscr{D}$ is the special distribution in (24); however, the argument for the general case is much the same.

Returning to Example 3, conditions (I) and (II) of Theorem 3 are easily checked: (I) holds near $x_0 = 0$, and (II) holds trivially since $\mathscr{D}(x) = span\{g(x)\} = span\{[0\ 1]^T\}$ is one-dimensional. In fact, in dimension $n = 2$, conditions (I) and (II) reduce to the single condition that

$$rank\left[g(x_0)[f, g](x_0)\right] = 2,$$

because then $g(x) \neq 0$ in some open set about $x_0$.

**Example 8.** The result of Example 7 showed that condition (II) of Theorem 3 does not hold for the system of Example 1. Let us check that condition (I) does hold: we just need to compute

$$ad_f^2 g(x) = \left[f, ad_f g\right](x) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} e^{-x_2} \\ x_1 \\ \frac{1}{2}x_1^2 \end{bmatrix} - \begin{bmatrix} 0 & -e^{-x_2} & 0 \\ 1 & 0 & 0 \\ x_1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ x_1 \end{bmatrix};$$

then, using the calculations from Example 7, we have

$$\left[g(x) \quad ad_f g(x) \quad ad_f^2 g(x)\right] = \begin{bmatrix} 1 & 0 & e^{-x_2} \\ 0 & 1 & 0 \\ 0 & x_1 & e^{-x_2} \end{bmatrix},$$

so the rank condition (I) is satisfied at every point.

We can now discuss condition (I) of Theorem 3 as a type of controllability condition known as a *local reachability* condition.

Suppose $f(x_0) = 0$, so that $x_0$ is an equilibrium for $x' = f(x)$. The rank condition (I) is exactly the condition that the local linearization defined by

$$x' = Ax + bu \equiv \frac{\partial f}{\partial x}(x_0)x + g(x_0)u \tag{30}$$

is completely controllable, that is, $rank[bAb \ldots A^{n-1}b] = n$. To see this, notice that if we write

$$f(x) = Ax + f_2(x), \quad g(x) = b + g_1(x),$$

where $(\partial f_2/\partial x)(x_0) = 0$ and $g_1(x_0) = 0$, then an induction proof shows that for each $k$,

$$ad_f^k g(x) = (-1)^k A^k b + p_k(x), \quad p_k(x_0) = 0. \tag{31}$$

Indeed, (31) holds when $k = 0$ because $g(x) = b + g_1(x)$ with $g_1(x_0) = 0$. The induction step follows by an appropriate bracket calculation using the expansions for $f$ and $g$. Thus, (I) is exactly the complete controllability condition for (30), provided $f(x_0) = 0$.

Let us say that a system $x' = f(x) + g(x)u$ is *locally reachable* at $x_0$ if there is an open set $U$ about $x_0$ such that every point $x \in U$ can be reached from $x_0$ in finite time by means of a control $u(t)$. Under the conditions of Theorem 3, the nonlinear system is locally reachable at $x_0$ because the system is locally equivalent to the completely controllable linear system (12), via the coordinate mapping $z = T(x)$ for $x \in U$ (and $z \in T(U)$), and the feedback transformation $u = \alpha(x) + \beta(x)v = \alpha(T^{-1}z) + \beta(T^{-1}z)v$. For if a control $v(t)$ transfers $z_0 = 0$ to the point $z_f$ in time $t_f$ while the trajectory $z(t)$ remains in $T(U)$ (and this can be done for system (12)), then the corresponding $u(t)$ keeps $x(t) = T^{-1}(z(t))$ within $U$. Therefore, transfers from $x_0$ to any $x_f$ in $U$ can be accomplished in finite time. If only the rank condition (I) holds, however, then provided $f(x_0) = 0$, local reachability of the nonlinear system at $x_0$ can still be proved using the controllability rank condition for the local linearization (30) together with the help of the Inverse Function Theorem. For one result of this type, which implies local reachability at $x_0$, see [**8**, Proposition 3.3, pp. 74–75].

The method of input-state linearization has been successful in addressing specific control problems in the areas of aircraft flight control and robotics [**10**, p. 207]. Here is a final example concerning the equations for a single-link robotic manipulator.

**Example 9.** [**6**, p. 528] [**10**, p. 242] The dynamical equations for a single-link, flexible-joint mechanism with negligible damping are

$$\begin{aligned} Iq_1'' + MGL\sin q_1 + k(q_1 - q_2) &= 0 \\ Jq_2'' - k(q_1 - q_2) &= u, \end{aligned}$$

where $q_1$ and $q_2$ are angular positions, $I$ and $J$ are moments of inertia, $k$ is a spring constant, $M$ is a mass, $G$ is the gravitational constant, $L$ is a distance, and $u$ is a motor torque input. By writing $x = [q_1\ q_1'\ q_2\ q_2']^T$, the four-dimensional state equations can be written as

$$x' = f(x) + g(x)u \equiv \begin{bmatrix} x_2 \\ -a\sin x_1 - b(x_1 - x_3) \\ x_4 \\ c(x_1 - x_3) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ d \end{bmatrix} u, \tag{32}$$

using these positive constants: $a = (MGL)/I$, $b = k/I$, $c = k/J$, and $d = 1/J$. The unforced system has an equilibrium at $x = 0$. To determine if this system is input-state linearizable near the origin, we check conditions (I), (II) of Theorem 3. First, compute

$$\frac{\partial f}{\partial x} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -a\cos x_1 - b & 0 & b & 0 \\ 0 & 0 & 0 & 1 \\ c & 0 & -c & 0 \end{bmatrix}.$$

In this case, the spanning vector fields for $\mathscr{D}$ are constant, and appropriate bracket calculations lead to the matrix needed in condition (I):

$$\begin{bmatrix} g & ad_f g & ad_f^2 g & ad_f^3 g \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & -bd \\ 0 & 0 & bd & 0 \\ 0 & -d & 0 & cd \\ d & 0 & -cd & 0 \end{bmatrix}.$$

Since this matrix has rank 4, condition (I) holds everywhere. It is immediate that the involutivity condition (II) holds everywhere, since $\mathscr{D}(x) = span\{g, ad_f g, ad_f^2 g\}$

is spanned by constant vector fields. By Theorem 3, system (32) is input-state linearizable, and the construction of the function $T_1$ with relative degree 4 can now be attempted. In fact, since we know that the null space of $dT_1(x)$ must be the flat distribution (linear subspace) defined by $span\{g, ad_f g, ad_f^2 g\}$, we take $T_1$ to be a linear function of $x_1$ alone: $T_1(x) = x_1$. The complete coordinate transformation $T(x)$ is then obtained from (16a), yielding

$$z = T(x) = \begin{bmatrix} T_1(x) \\ T_2(x) \\ T_3(x) \\ T_4(x) \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ -a \sin x_1 - b(x_1 - x_3) \\ -a x_2 \cos x_1 - b(x_2 - x_4) \end{bmatrix}.$$

Alternatively, one can construct $T_1$ directly from (16) for this example: the details appear in [**6**, pp. 528–529]. By defining the feedback transformation $u = \alpha(x) + \beta(x)v$ according to (17), the equations for $z$ are given by (12); equivalently, $z_1^{(4)} = v$. A control $v$ can now be designed that makes the link position $x_1 = z_1$ track a prespecified trajectory.

Clearly, systems that are fully input-state linearizable are very special. Nevertheless, as shown by the analysis in [**4**, pp. 162 – 172], if a system is not fully input-state linearizable as in Theorem 3, but an output function with relative degree $r < n$ is known, then the system can still be transformed to a partially linear normal form, which is quite useful in many control problems. As we noted in Example 3, the case when $r < n$ requires the examination of unobservable dynamics (usually called *zero dynamics*) [**4**, pp. 163–164]. However, the circle of ideas discussed here can be applied to a wide range of problems beyond the special conditions of Theorem 3.

**9. FURTHER READING.** The equivalence problem discussed here was first considered in [**1**]. For the origin of the use of Lie brackets in the study of reachability problems see [**3**] and the references therein. See also [**5**, pp. 1–2] for some historical insight on the introduction into control theory of differential-geometric ideas centered around the Lie bracket. Reference [**4**] presents important control methods using extensions of the basic ideas discussed in this article, and includes a development of the required differential-geometric concepts. References [**4**] and [**10**] discuss the important issue of stability of unobservable dynamics (called *zero dynamics*) that arose in Example 3, which comes from [**6**]. Information on controllability, observability, and numerous other issues appears in [**8**] and [**11**]. Reference [**8**, p. 59] contains a statement of the classical Frobenius' Theorem on integrability of a system of linear, first-order partial differential equations. See [**6**] and [**7**] for some engineering emphasis combined with excellent theoretical exposition. Much of the differential-geometric nonlinear control theory generalizes the geometric approach to linear control theory presented in [**13**]. Many references to the primary mathematical control literature may be found in the references listed.

REFERENCES


1. R. W. Brockett, Feedback invariants for nonlinear systems, in *Proceedings International Federation on Automatic Control Congress*, Helsinki, 1978, pp. 1115–1120.
2. J. K. Hale, *Ordinary Differential Equations*, Second Edition, Krieger, Malabar, 1980.
3. R. Hermann, "On the accessibility problem in control theory," in *International Symposium on Nonlinear Differential Equations and Nonlinear Mechanics* (eds. J. P. Lasalle, S. Lefschetz), Academic Press, New York, 1963, pp. 325–332.

4. A. Isidori, *Nonlinear Control Systems*, Third Edition, Springer, New York, 1995.
5. V. Jurdjevic, *Geometric Control Theory*, Cambridge Studies in Advanced Mathematics 51, Cambridge University Press, Cambridge, 1997.
6. H. K. Khalil, *Nonlinear Systems*, Second Edition, Prentice-Hall, Inc., Englewood Cliffs, 1996.
7. R. M. Murray, Z. Li, and S. S. Sastry, *A Mathematical Introduction to Robotic Manipulation*, CRC Press, Boca Raton, 1994.
8. H. Nijmeijer and A. J. van der Schaft, *Nonlinear Dynamical Control Systems*, Springer-Verlag, New York, 1990.
9. W. Rudin, *Principles of Mathematical Analysis*, Second Edition, McGraw-Hill, New York, 1964.
10. J. E. Slotine and W. Li, *Applied Nonlinear Control*, Prentice-Hall, Inc., Englewood Cliffs, 1991.
11. E. D. Sontag, *Mathematical Control Theory: Deterministic Finite Dimensional Systems*, Second Edition, Springer-Verlag, New York, 1998.
12. W. J. Terrell, Some Fundamental Control Theory I: Controllability, Observability, and Duality, *Amer. Math. Monthly*, to appear.
13. W. M. Wonham, *Linear Multivariable Control*, Third Edition, Springer-Verlag, New York, 1985.

**WILLIAM J. TERRELL** received his Ph.D. in Applied Mathematics from North Carolina State University in 1990. He is an Associate Professor in the Department of Mathematical Sciences at Virginia Commonwealth University, Richmond, Virginia. His research interests are in the area of differential equations with special focus on systems and control theory.
*Department of Mathematical Sciences, Virginia Commonwealth University, Richmond, VA 23284-2014*
*wterrell@atlas.vcu.edu*

## MATHEMATICS

C. O. Oakley, Haverford College

Mathematics is one component of any plan for liberal education. Mother of all the sciences, it is a builder of the imagination, a weaver of patterns of sheer thought, an intuitive dreamer, a poet. The study of mathematics cannot be replaced by any other activity that will train and develop man's purely logical faculties to the same level of rationality. Through countless dimensions, riding high the winds of intellectual adventure and filled with the zest of discovery, the mathematician tracks the heavens for harmony and eternal verity. There is not wholly unexpected surprise, but surprise nevertheless, that mathematics has direct application to the physical world about us. For mathematics, in a wilderness of tragedy and change, is a creature of the mind, born to the cry of humanity in search of an invariant reality, immutable in substance, unalterable with time. Mathematics is an infinity of flexibles forcing pure thought into a cosmos. It is an arc of austerity cutting realms of reason with geodesic grandeur. Mathematics is crystallized clarity, precision personified, beauty distilled and rigorously sublimated. The life of the spirit is a life of thought; the ideal of thought is truth; everlasting truth is the goal of mathematics.

MONTHLY 56 (1949) 19